Homework Assignment 2

Rushabh Ashok Dharia
CSCI-B555 Machine Learning

February 27, 2019

Note: Homework was discussed with Varun Miranda and Bhushan Malgaonkar

Ans 1.

$$f(x/\alpha) = \alpha^x (1-\alpha)^{1-x}$$

$$\text{Given } \alpha \text{ is selected from a uniform distribution(0,1)}$$

$$\text{Therefore } f(\alpha) = \frac{1}{1-0} = 1$$

$$argmax_\alpha(f(\alpha)f(x/\alpha)) = 1.\prod_{i=1}^{n} \alpha^x (1-\alpha)^{1-x}$$

$$\text{Taking log}$$

$$= \sum_{i=1}^{n} (x\log(\alpha) + (1-x)\log(1-\alpha))$$

$$\text{Taking derivative w.r.t } \alpha$$

$$0 = \frac{1}{\alpha}\sum_{i=1}^{n} x - \frac{1}{1-\alpha}\sum_{i=1}^{n}(1-x)$$

$$\frac{1}{\alpha}\sum_{i=1}^{n} x = \frac{1}{1-\alpha}\sum_{i=1}^{n}(1-x)$$

$$\sum_{i=1}^{n} x - \alpha\sum_{i=1}^{n} x = N\alpha - \alpha\sum_{i=1}^{n} x$$

$$\hat{\alpha}_{MAP} = \frac{1}{N}\sum_{i=1}^{n} x = \bar{x}$$

Ans 2.

$$argmax_{\theta} \prod_{i=0}^{n} \theta e^{-\theta x}$$

Taking log

$$\sum_{i=0}^{n} \log \theta - \theta x$$

Taking derivative w.r.t.θ

$$\sum_{i=0}^{n} (\frac{1}{\theta} - x_i) = 0$$

$$\frac{n}{\theta} = \sum_{i=0}^{n} x_i$$

$$\hat{\theta}_{MLE} = \frac{1}{\bar{x}}$$

Ans 3. *a).*

For H0,

$$f_X(x) = f_N(\frac{x+s}{a})$$

For H1,

$$f_X(x) = f_N(\frac{x-s}{b})$$

$$f_N(\frac{x+s}{a}) \underset{H1}{\overset{H0}{\gtrless}} f_N(\frac{x-s}{b})$$

Since $a = b$

$$f_N(\frac{x+s}{a}) \underset{H1}{\overset{H0}{\gtrless}} f_N(\frac{x-s}{a})$$

$$\frac{1}{\pi(1 + (\frac{x+s}{a})^2)} \underset{H1}{\overset{H0}{\gtrless}} \frac{1}{\pi(1 + (\frac{x-s}{a})^2)}$$

$$1 + (\frac{x-s}{a})^2 \underset{H1}{\overset{H0}{\gtrless}} 1 + (\frac{x+s}{a})^2$$

$$x^2 - 2xs + s^2 \underset{H1}{\overset{H0}{\gtrless}} x^2 + 2xs + s^2$$

$$0 \underset{H1}{\overset{H0}{\gtrless}} 4xs$$

Since s>0, s is positive

$$0 \underset{H1}{\overset{H0}{\gtrless}} x$$

$\square$

b).

$$r(\hat{y} = \int_{H0} f(x|y = H1).P(y = H1)dx + \int_{H1} f(x|y = H0).P(y = H0)dx$$

$$\text{Since priors are equal } P(y = H1) = P(y = H0) = \frac{1}{2}$$

$$r(\hat{y} = \frac{1}{2}\int_{H0} f(x|y = H1)dx + \int_{H1} f(x|y = H0)dx$$

$$= \frac{1}{2}\int_{-\infty}^{0} f_N(\frac{x+s}{a}) + \frac{1}{2}\int_{0}^{\infty} f_N(\frac{x-s}{b})$$

$$= \frac{1}{2}\int_{-\infty}^{0} \frac{1}{\pi(1+(\frac{x+s}{a})^2)} + \frac{1}{2}\int_{0}^{\infty} \frac{1}{\pi(1+(\frac{x-s}{b})^2)}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{0} \frac{1}{1+(\frac{x+s}{a})^2} + \frac{1}{2\pi}\int_{0}^{\infty} \frac{1}{1+(\frac{x-s}{b})^2}$$

$$\text{put } t = \frac{x+s}{a} \text{ and } u = \frac{x-s}{b}$$

When $x = -\infty$, $t = -\infty$. When $x = 0$, $t = \frac{s}{a}$

When $x = 0$, $u = \frac{-s}{b}$. When $x = \infty$, $u = \infty$

$$r(\hat{y}) = \frac{1}{2\pi}\int_{-\infty}^{s/a} \frac{1}{1+t^2}dt + \frac{1}{2\pi}\int_{-s/b}^{\infty} \frac{1}{1+u^2}du$$

$$= \frac{1}{2\pi}(\tan^{-1}(\frac{s}{a}) + \frac{\pi}{2} + \frac{\pi}{2} - \tan^{-1}(\frac{-s}{b}))$$

$$= \frac{1}{2} + \frac{1}{2\pi}(\tan^{-1}(\frac{s}{a}) - \tan^{-1}(\frac{-s}{b}))$$

$$= \frac{1}{2} + \frac{1}{2\pi}(\tan^{-1}(\frac{s}{a}) + \tan^{-1}(\frac{s}{b}))$$

(If $a == b$)

$$r(\hat{y}) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(\frac{s}{a})$$

$\square$

Ans 4. *a).*

$$
\begin{aligned}
f_{Y|X}(y|x) &= \frac{f_{X|Y}(x|y).f_Y(y)}{f_X(x)} \\
f_Y(y) &= \alpha e^{-y\alpha} \\
f_X(x) &= \int_0^\infty f_{X|Y}(x|y).f_Y(y)dy \\
&= \int_0^\infty ye^{-yx}.\alpha e^{-y\alpha}dy \\
&= \alpha \int_0^\infty ye^{-y(x+\alpha)}dy \\
&\text{Using Integration by parts and LIATE Rule} \\
&= \alpha(y\int e^{-y(x+\alpha)})dy - \int(\frac{d(y)}{dy}\int e^{-y(x+\alpha)})dy)dy) \\
&= \alpha[\frac{-ye^{-y(x+\alpha)}}{x+a} - \frac{e^{-y(x+\alpha)}}{(x+\alpha)^2}]_0^\infty \\
&= \alpha[\frac{-e^{-y(x+\alpha)}}{x+a}(y+\frac{1}{x+\alpha})]_0^\infty \\
&= \alpha[\frac{e^{-y(x+\alpha)}}{x+a}(y+\frac{1}{x+\alpha})]_\infty^0 \\
&= \alpha[\frac{1}{x+\alpha}(0+\frac{1}{x+\alpha}) - 0] \\
&= \frac{\alpha}{(x+\alpha)^2} \\
f_{Y|X}(y|x) &= \frac{ye^{-yx}\alpha e^{-y\alpha}}{\frac{\alpha}{(x+\alpha)^2}} \\
&= (x+\alpha)^2.y.e^{-y(x+\alpha)}
\end{aligned}
$$

□

b).

$$
\begin{aligned}
\hat{y}_{MAP} &= argmax_y f_Y(y).f_{X|Y}(x|y) \\
&= \alpha e^{-\alpha y}ye^{-yx} \\
&\text{Taking log} \\
&= log(\alpha) - \alpha y + log(y) - yx
\end{aligned}
$$

Taking derivative w.r.t. y

$$
\begin{aligned}
0 &= 0 - \alpha + \frac{1}{y} - x \\
\hat{y}_{MAP} &= \frac{1}{x+\alpha}
\end{aligned}
$$

□

c).

$$
\begin{aligned}
\hat{y}_{MMSE} &= E[y|x] \\
&= \int_0^\infty y f_{Y|X}(y|x) dy \\
&= \int_0^\infty y^2 e^{-(\alpha+x)y} (\alpha+x)^2 dy \\
&= (\alpha+x)^2 \int_0^\infty y^2 e^{-(\alpha+x)y} dy \\
&\text{Using integration by parts and LIATE Rule} \\
&= (\alpha+x)^2 (y^2 \int_0^\infty e^{-(\alpha+x)y} dy - \int_0^\infty (\frac{d}{dy} y^2 \int_0^\infty e^{-(\alpha+x)y} dy) dy) \\
&= (\alpha+x)^2 [(y^2 \frac{e^{-(\alpha+x)y}}{-(a+x)})_0^\infty + \frac{2}{a+x} \int_0^\infty y e^{-(\alpha+x)y} dy \\
&[\text{From a}) \int_0^\infty y e^{-(\alpha+x)y} dy = \frac{1}{(x+\alpha)^2}] \\
\hat{y}_{MMSE} &= (\alpha+x)^2 [-0 + 0 + \frac{2}{\alpha+x} . \frac{1}{(\alpha+x)^2}] \\
&= \frac{2}{\alpha+x}
\end{aligned}
$$

□

Ans 5. *a).*

$$
\begin{aligned}
P(S=1|V=1) &= P(S=1) \text{ (Since S is independent of V)} \\
P(S=1) &= P(S=1|G=0).P(G=0) + P(S=1|G=1).P(G=1) \\
&= (1-\gamma)\alpha + (1-\beta)(1-\alpha)
\end{aligned}
$$

□

b).

$$
\begin{aligned}
P(S=1|V=0) &= P(S=1) \text{ (Since S is independent of V)} \\
&= P(S=1|V=1) \\
&= P(S=1|G=0).P(G=0) + P(S=1|G=1).P(G=1) \\
&= (1-\gamma)\alpha + (1-\beta)(1-\alpha)
\end{aligned}
$$

Explanation : S is independent of V. Hence value of V doesn't affect S

□

Ans 6. *a).*

$$
\begin{aligned}
P(X_2 = salmon) &= P(X_1 = Winter).P(X_2 = salmon|X_1 = Winter) + \\
&\quad P(X_1 = Autumn).P(X_2 = salmon|X_1 = Autumn) \\
&= 0.5 * 0.9 + 0.5 * 0.8 \\
&= 0.85 \\
P(X_2 = seabass) &= P(X_1 = Winter).P(X_2 = seabass|X_1 = Winter) + \\
&\quad P(X_1 = Autumn).P(X_2 = seabass|X_1 = Autumn) \\
&= 0.5 * 0.1 + 0.5 * 0.2 \\
&= 0.15
\end{aligned}
$$

□

b).

$$
\begin{aligned}
P(Winter) &= P(X_1 = Winter) * P(X_2 = salmon|X_1 = Winter) * P(X_3 = Dark|X_2 = salmon) \\
&* P(X_4 = Wide|X_2 = salmon) + P(X_1 = Winter) * P(X_2 = seabass|X_1 = Winter) \\
&* P(X_3 = Dark|X_2 = seabass) * P(X_4 = Wide|X_2 = seasbass) \\
&= 0.25(0.9 * 0.34 * 0.4 + 0.1 * 0.1 * 0.95) \\
&= 0.0323 \\
P(Spring) &= P(X_1 = Spring) * P(X_2 = salmon|X_1 = Spring) * P(X_3 = Dark|X_2 = salmon) \\
&* P(X_4 = Wide|X_2 = salmon) + P(X_1 = Spring) * P(X_2 = seabass|X_1 = Spring) \\
&* P(X_3 = Dark|X_2 = seabass) * P(X_4 = Wide|X_2 = seasbass) \\
&= 0.25(0.3 * 0.34 * 0.4 + 0.7 * 0.1 * 0.95) \\
&= 0.0268 \\
P(Summer) &= P(X_1 = Summer) * P(X_2 = salmon|X_1 = Summer) * P(X_3 = Dark|X_2 = salmon) \\
&* P(X_4 = Wide|X_2 = salmon) + P(X_1 = Summer) * P(X_2 = seabass|X_1 = Summer) \\
&* P(X_3 = Dark|X_2 = seabass) * P(X_4 = Wide|X_2 = seasbass) \\
&= 0.25(0.4 * 0.34 * 0.4 + 0.6 * 0.1 * 0.95) \\
&= 0.02785 \\
P(Autumn) &= P(X_1 = Autumn) * P(X_2 = salmon|X_1 = Autumn) * P(X_3 = Dark|X_2 = salmon) \\
&* P(X_4 = Wide|X_2 = salmon) + P(X_1 = Autumn) * P(X_2 = seabass|X_1 = Autumn) \\
&* P(X_3 = Dark|X_2 = seabass) * P(X_4 = Wide|X_2 = seasbass) \\
&= 0.25(0.8 * 0.34 * 0.4 + 0.2 * 0.1 * 0.95) \\
&= 0.03195 \\
&\text{It is most likely to be winter}
\end{aligned}
$$

□

Ans 7. *a).*

$$\text{To show that, } E[x] = \sum_k \pi_k \mu_k$$

$$\text{Given, } p[x] = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)$$

$$E[x] = \sum_k \frac{\pi_k}{\sqrt{2\pi_k \sigma_k}} \int_{-\infty}^{\infty} x e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx$$

$$\text{Let, } x = x + \mu_k - \mu_k$$

$$= \sum_k \frac{\pi_k}{\sqrt{2\pi_k \sigma_k}} [\int_{-\infty}^{\infty} (x-\mu_k) e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx + \int_{-\infty}^{\infty} \mu_k e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx] \quad -[1]$$

$$\text{For } \int_{-\infty}^{\infty} (x-\mu_k) e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx$$

$$\text{Put } e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} = t$$

$$\frac{dt}{dx} = -\frac{x-\mu_k}{\sigma_k} . e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2}$$

$$-\sigma_k dt = (x-\mu_k) e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx$$

$$\int_{-\infty}^{\infty} (x-\mu_k) e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx = -\sigma_k \int_{-\infty}^{\infty} dt$$

$$= -\sigma_k [t]_{-\infty}^{\infty}$$

$$= 0 \quad -[2]$$

$$\text{Now } \int_{-\infty}^{\infty} \mu_k e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx = \mu_k \int_{-\infty}^{\infty} e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx$$

$$= \mu_k . \sqrt{2\pi_k \sigma_k} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi_k \sigma_k}} e^{\frac{-1}{2}(\frac{x-\mu_k}{\sigma_k})^2} dx$$

$$= \mu_k . \sqrt{2\pi_k \sigma_k}(1) \quad -[3]$$

Substituting [2] and [3] in [1]

$$E[x] = \Sigma_k \frac{\pi_k}{\sqrt{2\pi_k \sigma_k}} (0 + \mu_k . \sqrt{2\pi_k \sigma_k}(1))$$

$$= \Sigma_k \pi_k \mu_k$$

□

b).

$$\begin{aligned}
cov[x] &= cov[x, x^T] \\
&= E[(x-\mu_x)(x-\mu_x)^T] \\
&= E[xx^T - x\mu^T - \mu x^T - \mu\mu^T] \\
&= E[xx^T] - \mu_x^T E[x] - \mu_x E[x]^T + \mu_x\mu_x^T \\
&= E[xx^T] - \mu_x^T\mu_x - \mu_x\mu_x^T + \mu\mu^T \\
\text{Therefore, } cov[x] &= \Sigma = E[xx^T] + \mu_x\mu_x^T \\
\text{Therefore, } E[xx^T] &= \Sigma - \mu_x\mu_x^T \\
\text{For k Gaussians, } E[x] &= \Sigma_k\pi_k\mu_k \\
\text{Therefore, } E[XX^T] &= \Sigma_k\pi_k(\Sigma_k - \mu_x\mu_x^T) \\
\text{Finally, } cov[x] &= E[xx^T] - E[x]E[x]^T \\
\text{Therefore, } cov[x] &= \Sigma_k\pi_k(\Sigma_k - \mu_x\mu_x^T) - E[x]E[x]^T
\end{aligned}$$

□

Ans 8. b)
As shown in the plots notebook attached σ decreases as number of Gaussians increases.
This is because the probability of a point lying in each Gaussian decreases with increase in number of G.
Also, we can see that 2 Gaussians in k=4 almost overlap as we are trying to over fit the GMM.
Also, the $P(c_i)$ for each Gaussian decreases with increase in number of Gaussians